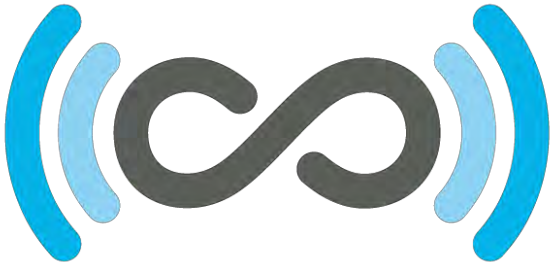


Open Data for Research Analytics: The SHARE Notification Service

Andrew Sallans, Center for Open Science

NORDP Annual Research Development Conference

Spring 2015

Founded by Academic Leaders
Built with Open Technology

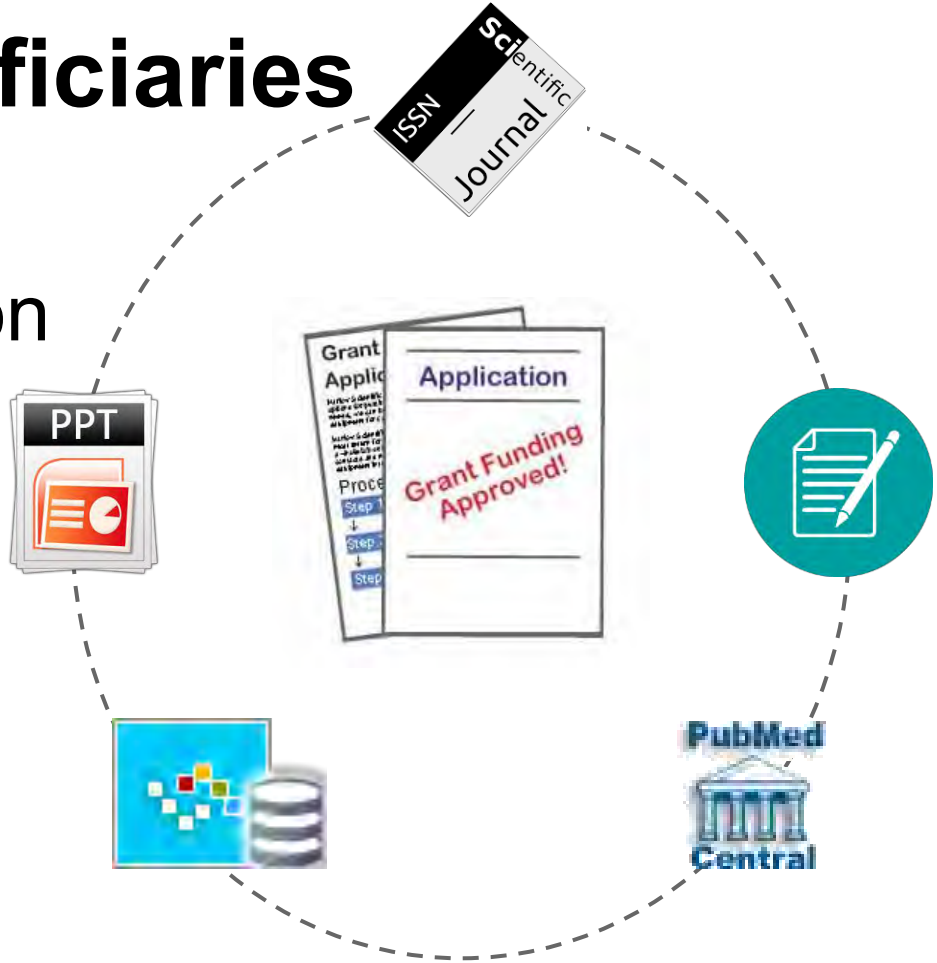
Research universities are long-lived and are mission-driven to generate, make accessible, and preserve over time new knowledge and understanding.

What is SHARE?

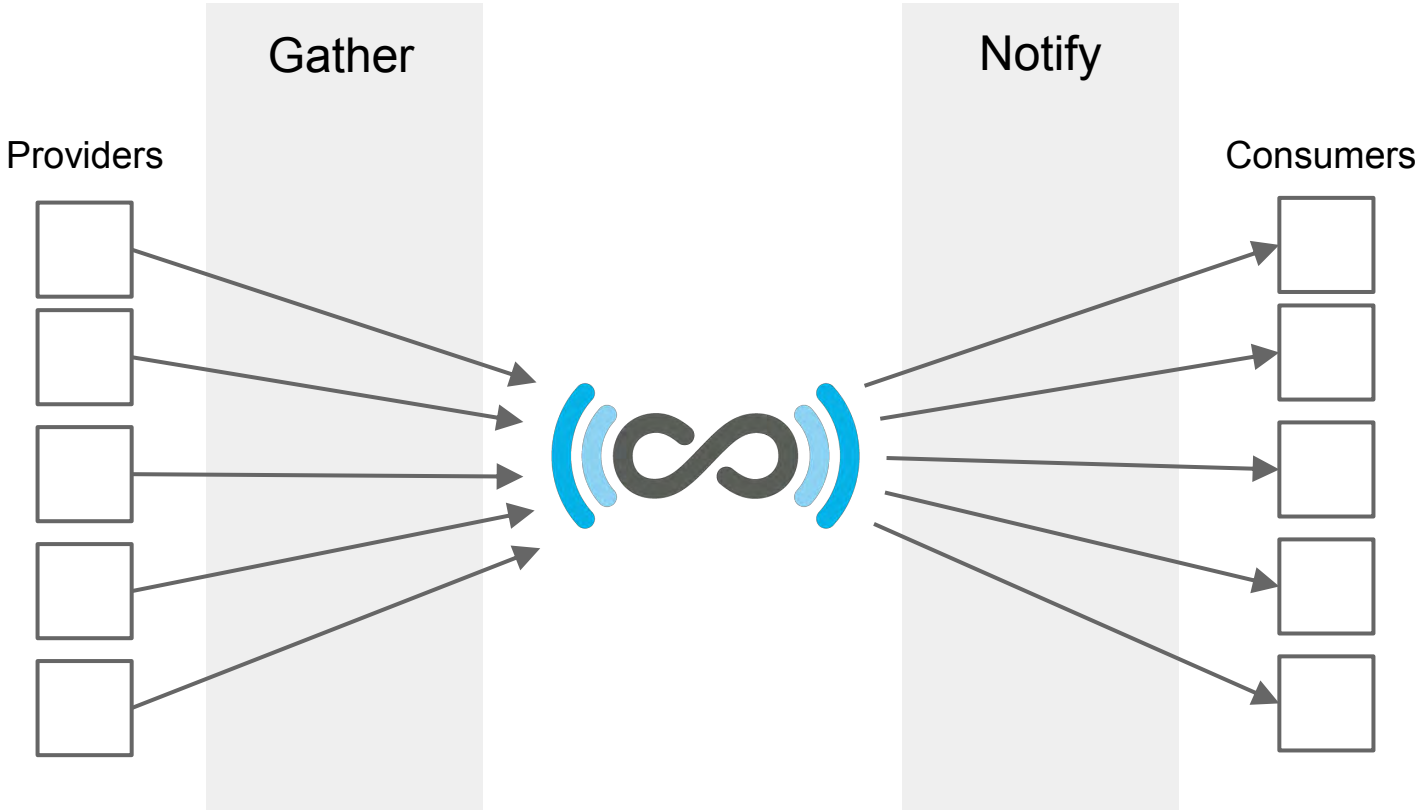
SHARE is creating an open data set about research activities across their life cycle. This comprehensive inventory of research will be widely accessible, discoverable, and reusable.

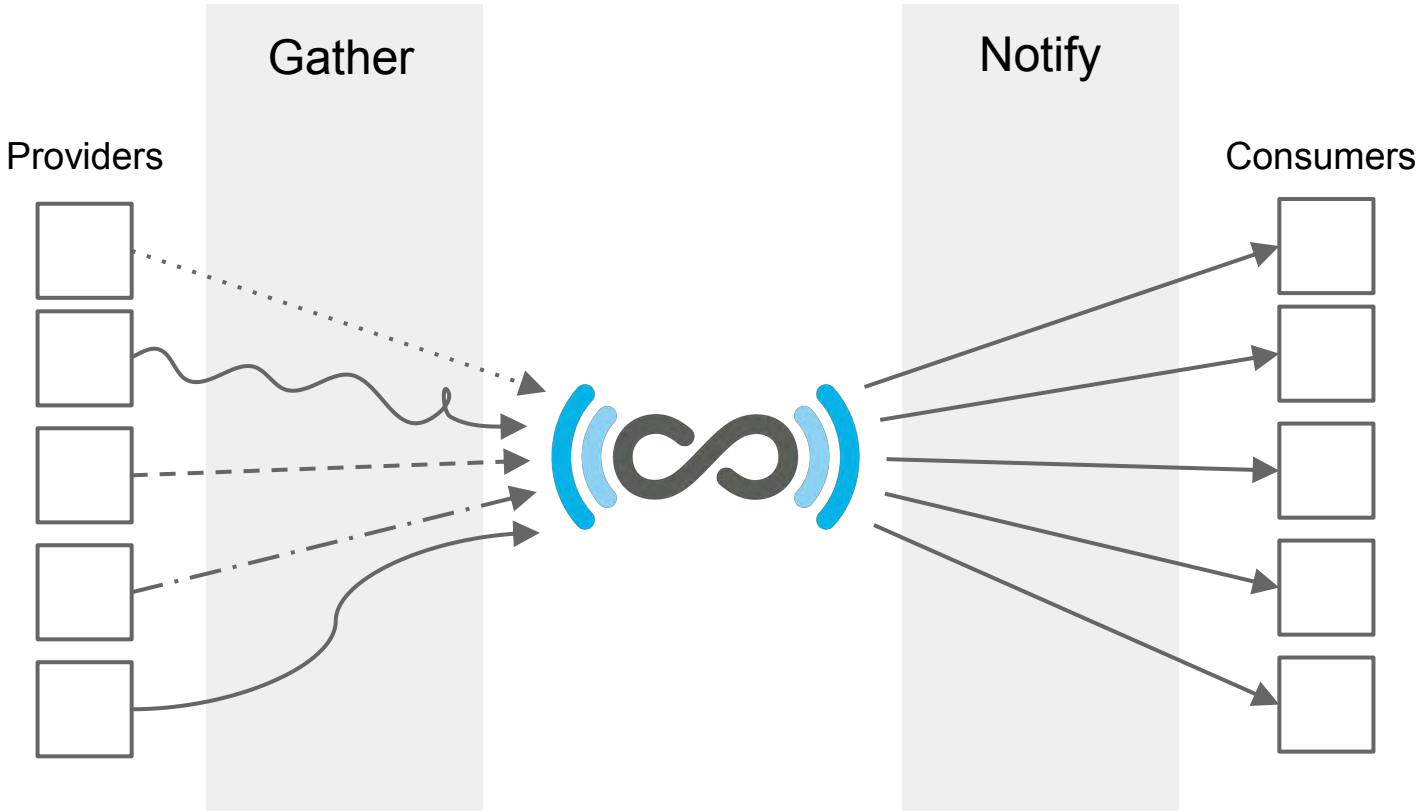
Stakeholders/beneficiaries

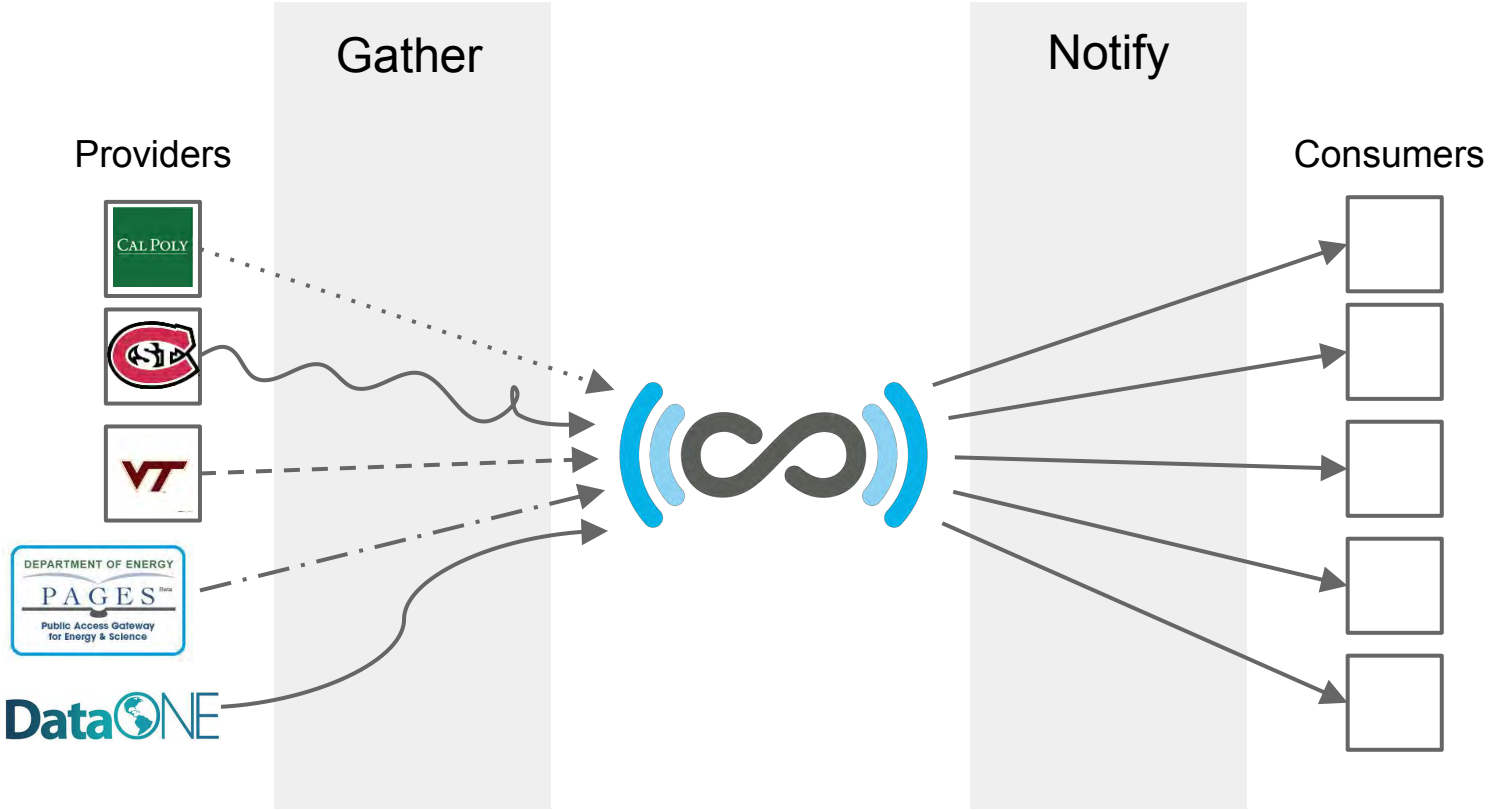
- Researchers
- Higher ed administration
- Libraries
- Repositories
- Research funders



Open data allows for innovation


Gather
Notify
Providers
Consumers


Gather
Notify
Providers
Consumers


Gather
Notify
Providers
Consumers
CAL POLY
DEPARTMENT OF ENERGY
PAGES
Public Access Gateway for Energy & Science
DataONE

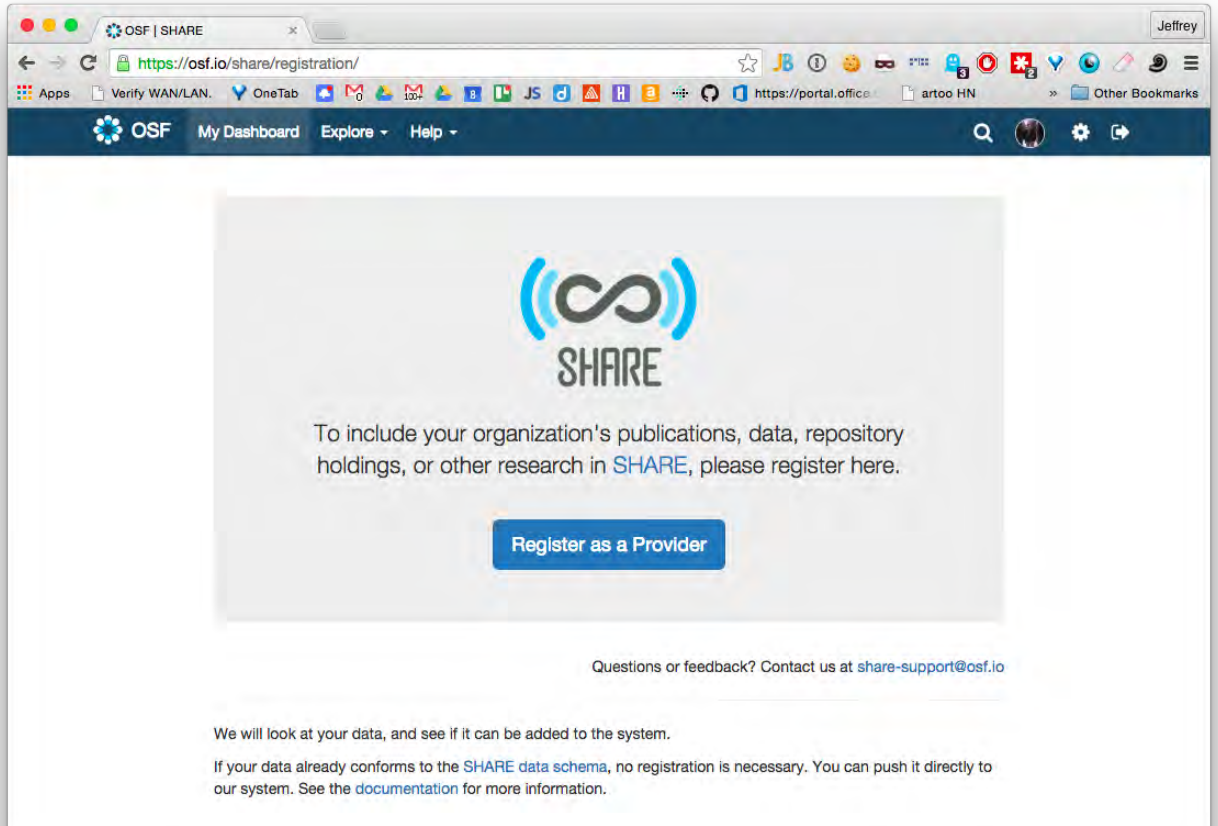
Scholars Portal Dataverse
A Service of the Ontario Council of University Libraries
Valparaiso University
arXiv.org
ClinicalTrials.gov
A service of the U.S. National Institutes of Health
BioMed Central
ASU
Arizona State University
tDAR
IDEALS
UT DR
University of Texas
Digital Repository
figshare
Carnegie Mellon University
University Libraries
University of Washington
DataONE
Texas State University
The rising STAR of Texas
Penn Libraries
University of Pennsylvania
eScholarship
University of California
PLoS
Public Library of Science
Trinity University
DSpace@MIT
Department of Energy
PAGES
Public Access Gateway for Energy & Science
VT
CSU
SIU Southern Illinois University
Carbondale
PubMed
DMPTool
Columbia University
Academic Commons
Wayne State
Cal Poly
SciTech Connect
crossref

After only a few months…

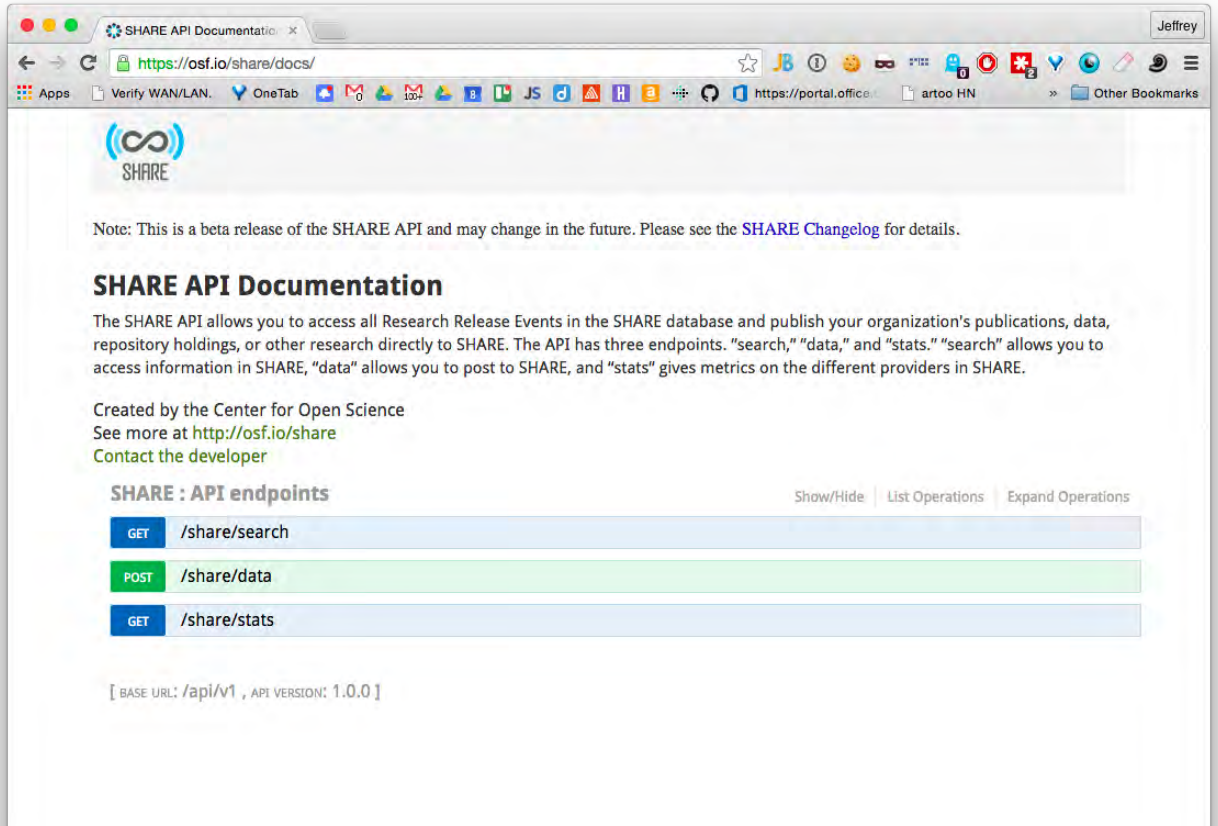
700,000+ research release events

30 content providers
(about 10 more in development)

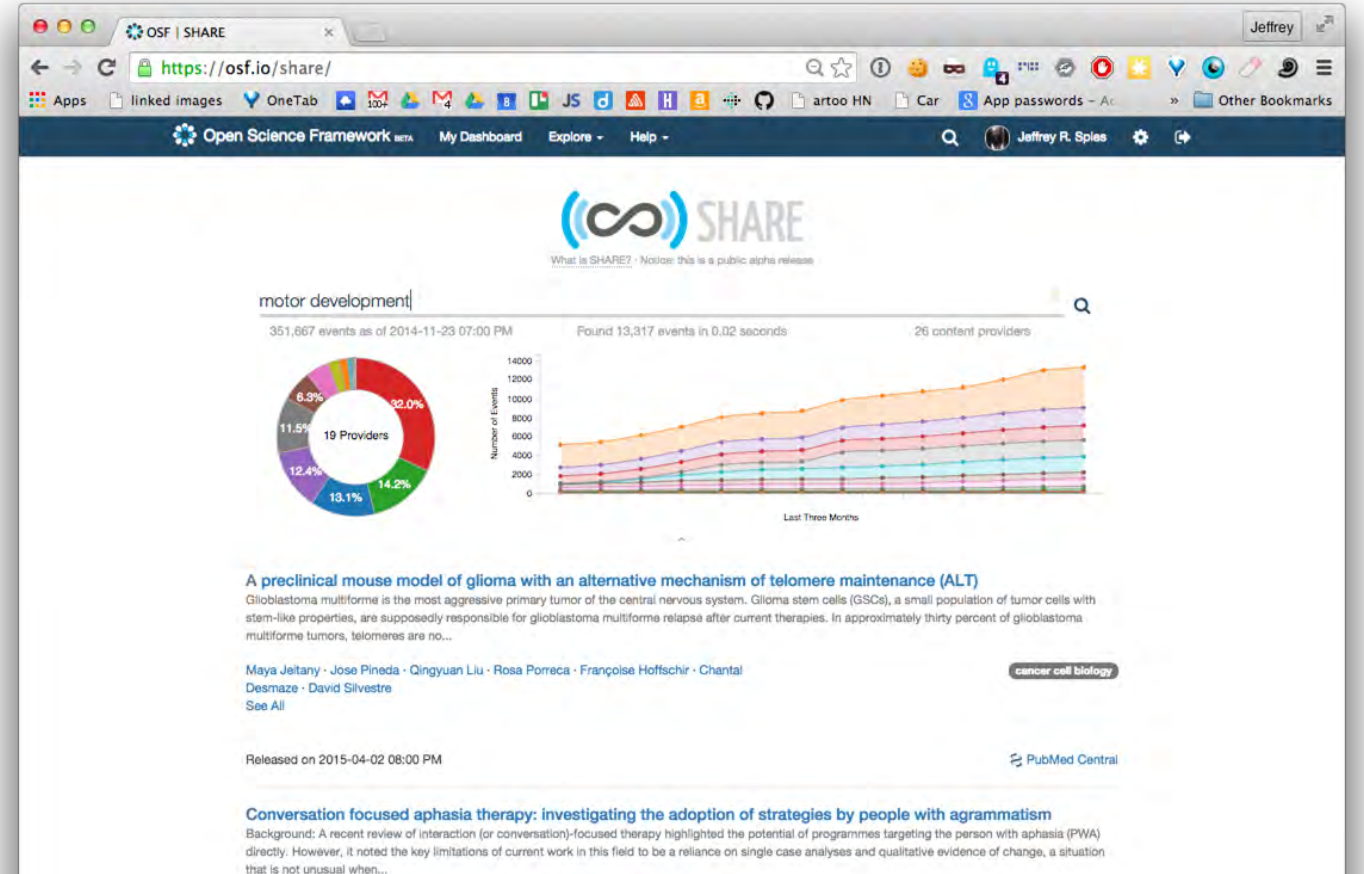
osf.io/share/registration



osf.io/share/docs



http://osf.io/share



What does this mean for you?

- Competitive analyses
- Matchmaking
- Broad scope of content

- FREE TO USE
- DATA IS OPEN
- INVITATION TO CONTRIBUTE

Under the hood

- Modularity
- Reliability
- Flexibility
- Scalability

python
TM

django

cassandra

fluentd

TOKUMX

kibana

Sentry

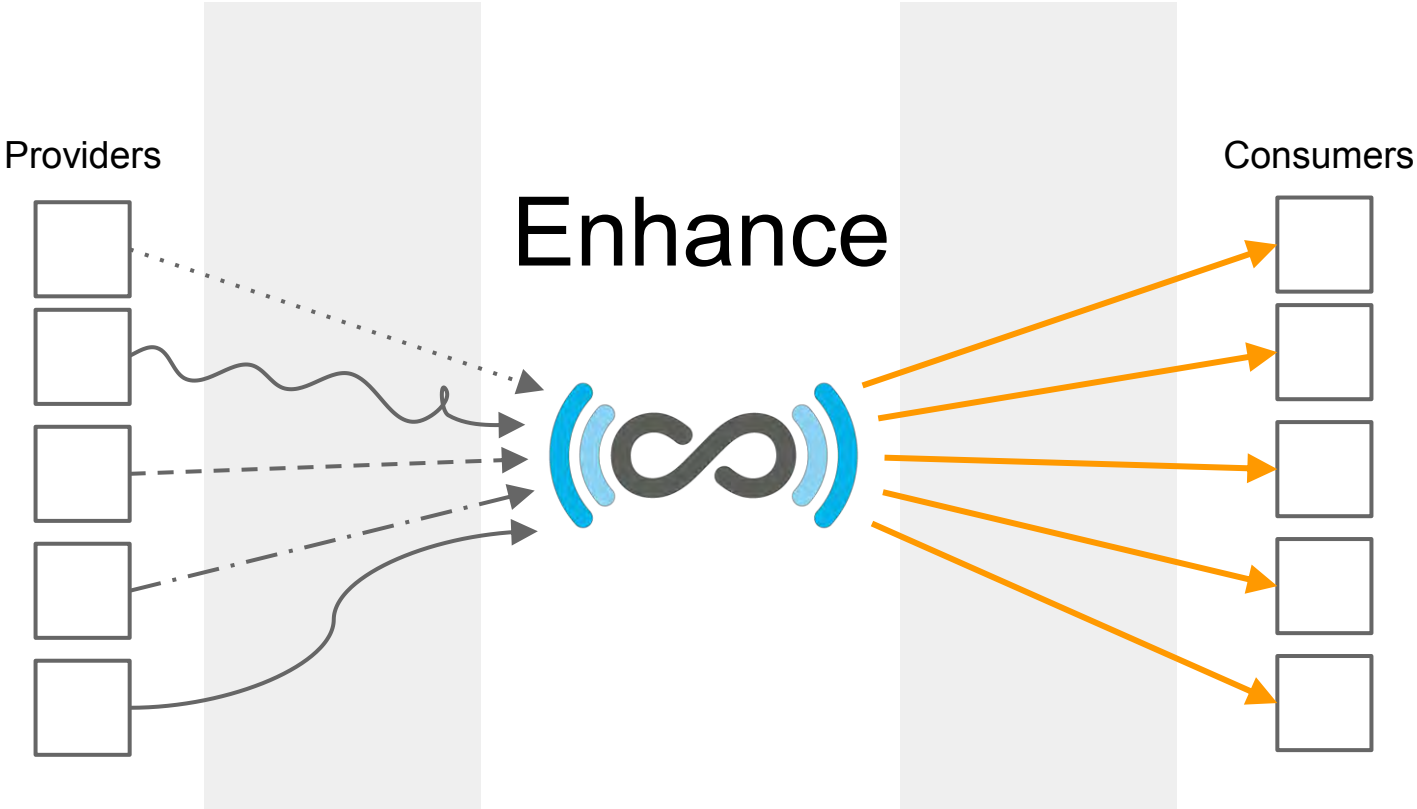
Live demo

What we've learned

- Metadata is used inconsistently across the research ecosystem
- Providers are unsure of their rights to share metadata
- Solutions will involve improvements to workflow, policy, and technical infrastructure

Where we're going

- Analyze stakeholders' current processes through which research passes from conception to completion
- Promote and reward policies that encourage data sharing
- Deploy multiple strategies to enhance metadata in open data set


Providers
Enhance
Consumers

Enhancement

- Open
- Collect
- Identify
- Connect

SHARE is a Community Project

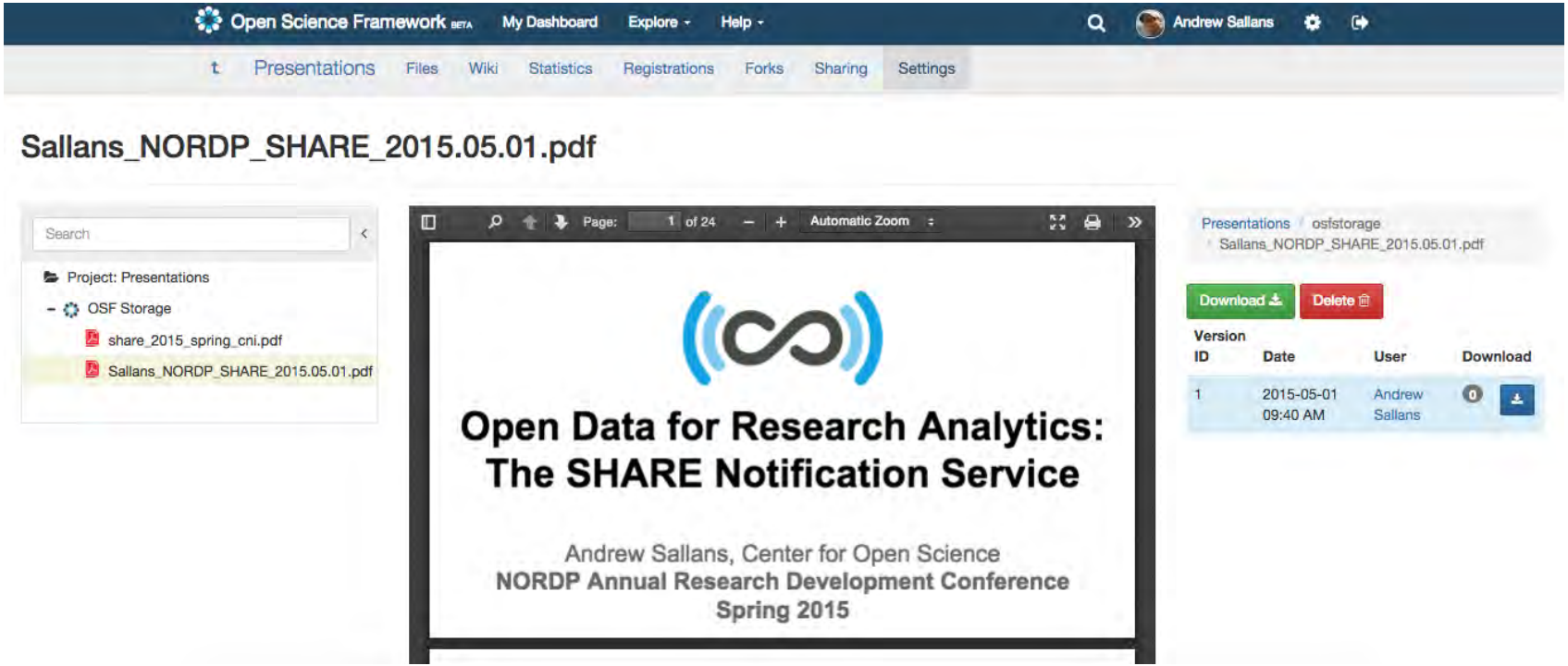
Opportunities for participation:

- **Infrastructure**
  https://github.com/centerforopenscience/share
  https://osf.io/wur56/wiki/

- **Policy/Workflow**
  SHARE Working Group
  Standards, identifiers, behaviors

Help us make this work for you

- How would you use SHARE?
- What would you need it to connect to?
- What content would it need to have?
- Interest in piloting?

Find this presentation at

https://osf.io/mn76t/